



November 21, 2024

VIA ECF

Honorable Ona T. Wang
 United States District Court
 500 Pearl Street
 New York, NY 10007

Re: *The New York Times Company v. Microsoft Corp., et al.*, No. 23-cv-11195-SHS; *Daily News, L P, et al. v. Microsoft Corp., et al.*, 1:24-cv-3285; *The Center for Investigative Reporting, Inc. v. OpenAI, Inc.*, 1:24-cv-04872: Response to Motion at ECF No. 309

Dear Magistrate Judge Wang:

Microsoft submits this response to the November 18, 2024 letter from Plaintiffs The New York Times Company (“The Times”), Daily News, LP et al., and The Center for Investigative Reporting, Inc. (collectively “News Plaintiffs”) regarding text and social media messages (ECF No. 309). While News Plaintiffs purport to seek to establish a “framework” governing collection and production of text messages, that is not what they have asked for, at least with respect to the two discrete issues that directly pertain to Microsoft. Instead, the relief sought by News Plaintiffs would require Microsoft to re-do collection and preparation for production of text messages and to undertake an inefficient collection of Slack messages on channels owned, controlled, and ultimately produced by OpenAI.

I. News Plaintiffs’ Demand for Production of Text Messages in 24-Hour Increments Is Arbitrary, Unnecessary, and Unduly Burdensome.

News Plaintiffs’ dispute over the format of text messages attempts to create a problem where none exists. As Microsoft communicated to The Times, Microsoft will produce text messages “in a usable format for plaintiffs, i.e., in sequential order with ascertainable dates. In other words, Microsoft will produce text messages in a manner that permits plaintiffs to conduct their necessary review without undue complication or burden.” ECF No. 309-2 at 3; *see also* ECF No. 309-3 at 3 (“[T]he text messages will be produced in a format that is usable for plaintiffs.”). News Plaintiffs say nothing about why this is not legally or practically sufficient, and if Plaintiffs wanted this specific format, they should have proposed it months ago during negotiation of the ESI protocol, before Microsoft had already begun to collect and review text messages. Instituting this arbitrary requirement now will require Microsoft to re-collect and/or manually re-format its existing collection for no purpose.¹

¹ Plaintiffs complain that “Microsoft refus[ed] to make a counterproposal.” ECF No. 309 at 3. The Times itself did not make a “proposal” or ask for Microsoft’s; it demanded that Microsoft “confirm” a format the parties never discussed. ECF No. 309-3 at 5-6.

Honorable Ona T. Wang

- 2 -

November 21, 2024

To put any questions to rest about the parade of horribles News Plaintiffs say will occur, *see* ECF No. 309 at 3 (“The 24-hour increment is important for . . . ”): Microsoft is producing text messages as single-page documents—what appears on one standard-sized page will be a single document. Where the single page does not provide the date(s) of the messages or does not “captur[e] relevant context,” one or more pages preceding and/or following that page will be produced. The pages will be produced in sequential order. Between inclusion of additional pages where appropriate and sequential production, the exchange will not be “hard[] to follow.” Nor will it “artificially inflate hit counts” by adding these single pages to hit counts.²

Simply put, there is nothing magic about a 24-hour increment; relevant portion(s) of a given message may be confined to a minute’s-worth of messages or it may stretch across days. What matters is that News Plaintiffs can review and understand the documents without undue burden. They will be able to.

II. Search and Collection of Slack Messages with OpenAI Employees on Slack Channels Controlled by OpenAI Would Have No Benefit to the Case, and the Burden to Microsoft Is Significant.

News Plaintiffs’ argument about Slack messages between OpenAI and Microsoft employees ignores the main point of what Microsoft communicated to Plaintiffs: these messages on Slack channels controlled by OpenAI are not easily collected by Microsoft. They are readily collected by OpenAI and indeed are already being searched for by OpenAI, making this discovery *entirely* cumulative and disproportionate.

As Microsoft communicated to Plaintiffs, the Slack channels at issue are controlled by and in the possession of OpenAI. ECF No. 309-3 at 3. That is, Microsoft does not have administrative or back-end access to these messages that would allow Microsoft to search or collect the messages efficiently. Microsoft would have to access each custodian’s device and manually capture each potentially relevant conversation, likely through screenshots.³ OpenAI, in contrast, as the owner and administrator of these channels, is able to export conversations in an efficient manner.

Microsoft understands that OpenAI has collected and is in the process of producing responsive Slack messages. Accordingly, the discovery Plaintiffs seek here is wildly inefficient and also non-sensical because it is entirely duplicative and provides no benefit to the case.

Instead of explaining how this discovery is proportionate, Plaintiffs state, with no basis, that Microsoft employees “used their own devices [for Slack messaging], which are the same ones Microsoft will be searching for Text Messages.” ECF No. 309 at 3. This short statement contains several unsupported and inaccurate assumptions, but the heart of the problem is that News Plaintiffs’ position ignores that accessing a device is only step one in collecting these messages.

² Hit counts for Plaintiffs’ requested search terms have repeatedly stretched into the millions without text messages.

³ Based on counsel’s investigation to date, Slack does not allow an individual user who is not an administrator or owner to easily download conversations.

Honorable Ona T. Wang

- 3 -

November 21, 2024

The Slack messages will still need to be screenshotted, one by one. It is not surprising that News Plaintiffs missed the point: instead of engaging in a cooperative process, Plaintiffs hustled this issue into an unnecessary dispute. After learning that Microsoft and OpenAI employees used Slack to communicate, News Plaintiffs rushed to filing a letter on this issue just two business days later. This rush to file resulted in the inaccurate and off-topic argument presented in its letter brief.

In short, News Plaintiffs will receive relevant and responsive Slack messages at issue from OpenAI, and requiring Microsoft to engage in duplicative collection of the very same Slack messages when it does not own or control the channels on which they reside makes no sense.

* * *

Taking a step back, these issues illustrate a larger discovery problem in these cases: instead of seeking to resolve issues, Plaintiffs are creating them. Instead of working cooperatively to make discovery as efficient as possible for all parties, Plaintiffs look for ways to slow down Defendants' progress in collecting and producing the documents that are actually relevant to this case.

There simply has never been any indication that text messages from Microsoft or Slack communications between Defendants will be a source of information for the core issues in this case—namely the contents of the training data underlying the models at issue and whether Defendants' activities constitute a fair use under the copyright laws. Information relevant to these core inquiries is just not going to be found in text messages, nor do Plaintiffs suggest that it could be. It makes no sense that text or Slack messages would make any of the core issues in dispute in this case any more or less likely.

Nevertheless, Microsoft has agreed to produce these messages, and the relevant Slack messages between Microsoft and OpenAI that are in OpenAI's possession on channels controlled by OpenAI will also be produced. There is nothing more here for the Court to do.

Respectfully submitted,

/s/ Annette L. Hurst

Annette L. Hurst

Respectfully submitted,

/s/ Jared B. Briant

Jared B. Briant

Counsel for Defendant Microsoft Corporation